Avoiding Single Point of Failure in Triple Play

"Physician, heal thyself"

Howard C. Berkowitz

hcb@netcases.net
What's the difference between used car salesmen and network service sales?
Premises for the Provider

• A sick physician cures nobody

• Service provider infrastructure has to be up before customers
  — Classic example: if your VoIP PBX catches on fire, who calls the fire department?
  — Observation 1: Whenever the magic smoke starts leaking out of a router, the router will soon fail
  — Observation 2: With optical communications, if the magic mirror breaks, there will be more than 7 years of missing lambdas
Premises for the Customer

- Accept provider prioritization
  - Polite: "In the unlikely event of a cabin depressurization, put on your mask before helping others"
  - Real: "In the unlikely event of a cabin depressurization, grab the mask, flying around in hurricane-force wind, and get it over your face before you pass out. You have seconds if it's real."

- Know the realities of disaster mitigation
  - You have your own responsibilities
  - Sometimes multiple providers are an answer
  - Sometimes a single provider that engineers high availability is the answer
What is the Problem to be Solved?

• You are a service provider
  — You offer Internet, video and telephony
  — Your customers look to you for disaster recovery

• *Disaster Recovery involves More than your Plant*
  — Customer sites, connectivity have to be prepared to work with your approach to disaster recovery
  — Multihomed customers may need to coordinate with several providers
    - Professional services opportunity for one provider to manage

• To the customer, what are the perceived needs?
Components of Disaster Recovery

- Knowing a disaster is coming or has arrived
- Communicating about it
  - Emergency responders
  - Vendors and support infrastructure
  - Staff
  - Customers
- Preparedness and response
  - Physical plant, power, etc.
  - Physical connectivity
  - IP-transparent failover
  - IP multihoming and failover
  - Transport/session response
  - Application response
What breaks in the General Network?

- Physical site
- Upstream connectivity
- User connectivity
- IP and sub-IP infrastructure
- Load distribution
- Applications
What breaks for VoIP?

- Customer premises codec/router
- Customer premises VoIP PBX
- Session Border Controllers
- Softswitch
  - Redundant Call Agents out of sync
  - Other common elements
- Upstream Connectivity
  - SIP
  - GR.303
- 3rd Party services: directory, E911, LNP
CA-CA synchronization in VoIP
What breaks for video?

- Download or other video feed
- Multicasting routers
- Insufficient multicast configuration
- Confused IGMP
  - switches not snooping
  - hosts not requesting
- Video-on-demand servers
- Access control for premium service
How do you find out about disasters?

- Mailing lists (open and closed)
- Emergency services
- News organizations (Radio & TV, Cable + outside antenna, in NOC)
- Looking outside (window not in critical areas)
How to talk about them

- Assume some aspects of telephone, SMS, pagers, email, etc., will be down or restricted in use
- If you can't talk to vendors, upstreams, emergency services, you will do your customers no good
- Still, you want to keep customers informed
Redundant Human Connectivity is Essential

- Telephone numbers not toll-free only
- Landline and cellular, possibly satellite
- If you can be considered critical infrastructure, obtain priority access
  - In US, National Communications System programs
    - GETS
    - WPS
Defining Service Expectations

- Availability
- "Classic SLA"
  - SLA for interactive applications
  - SLA for mission-critical data (computer-to-computer)
  - SLA for voice
  - SLA for video
    - TV
    - Videoconference
Availability Expectations: Your Site
Don't Forget Backup Power...
Power Supply is not trivial (1)

- Consider >1 utility feed
- Generators
  - Need fuel
    - Diesel has limited storage life
    - Typically test system weekly
    - Arrange for fuel deliveries
  - Other needs
    - Starting batteries
    - Air filters
    - Physical security
  - Placement in building
    - Rising water
    - Falling burning fuel
Power Supply is not trivial (2)

- Power feed transfer (utility, generator)
- UPS
  - Liquid electrolyte may have fire code restrictions on placement
  - Need connectivity to all sources
  - Redundancy here is important
    - Failover
    - Maintenance
Specifying Availability

Rules are always subject to interpretation

Ferengi Rule of Acquisition #284
Specifying Availability for Business Services (8/5)

• Period of coverage
  — Period of technical support availability if different
• Restrictions on offered load under disaster mode?
• Maintenance windows?
• When does an outage begin? end?
  — see quality discussion later in this presentation
• Opportunities for less-than-ideal backup?
• Pricing incentives?
Traffic Engineering (24/7): VoIP, TVoIP

- Throughput
- Need for consistent latency (minimize jitter)
- Availability
- Enough bandwidth
- Bandwidth in the right place
- Transient congestion avoidance
- Alternative ways to supply resources

- There will be single points of failure in the local loop
- Consider physical multihoming for SOHO
  You can probably get bulk rates on other media (e.g., DSL if you are cable or vice-versa)
Higher Layer Threats & Responses

• Single server failure or maintenance downtime
• Individual overloaded servers at single site
• Overloaded site or servers, but sufficient overall capacity
• Server crash

• Clustered servers at site; cold, warm, hot standby
• Local load distribution inside cluster
• Global load distribution among multiple clusters and sites
• Backups, checkpoints, mirroring
Lower Layer Threats & Responses

- Routing system failure
- Failure of direct provider or upstream links
- Failure of customer router on LAN
- Single medium failure between customer and ISP
- Multiple ISPs
- Multiple connection to single provider. Diversity contracts.
- VRRP/HSRP. BGP peering to loopbacks.
- Inverse multiplexing. SONET. Dial/ISDN backup. Local loop diversity
Sub-IP
There remain L2 switching roles

- Backup server farms with duplicate IP addressing on "outside"
  - "Public" server addresses could be anycast, or contained in multiple DNS entries
- Still need unique maintenance addresses
- Even with IP
  - Consider private VLAN in local distribution
- VRRP/HSRP between locations
- Again, unique maintenance addresses
It's not just IP

- IP is certainly your base for communications
- But you may use it in atypical ways
  - Anycast
  - L2 failover with duplicate but standby IP addresses
  - Virtual IP addresses (e.g., VRRP, HSRP)
VoIP Providers

- Need upstream PSTN connectivity
  - Traditional is GR.303
    - May need PWE3 for internal IP
    - PWE3 tells upstream you are PDH (DS1/3), SONET, ATM
  - SIP growing in utility
    - True SIP peering
    - Layer 3 connectivity to peers, SIP one aspect
    - New motivations for peering, exchange points
TV over IP

- **Multicast application**
  - Ignoring Pay-Per-View for now...
  - How many feeds per head end multicast router?
  - May need
    - Your own local video storage
    - Content distribution network

- **What is the physical plant?**
  - Copper coax won’t support extensive HDTV or IP
  - Need bidirectional
  - Optical sooner or later
    - FTTC? FTTB? FTTH?
Some Routing Scenarios

Registered address space
Provider 1
Provider 2

Registered or private address space

Private address space
Question:

What is the most important machine in the hospital?
Basic Machine Thoughts

- When putting redundant processors into a machine
  - Consider maintenance: can you update backup blade?
  - Do you need additional machine for hot update?
- When putting in redundant line cards
  - Separate interface processors when that’s the model
  - If there are multiple fabric/backplane busses, be sure to use different ones
Midboxes: who troubleshoots?

- Load Aware DNS
- Load Sharing NAT
- PAT/NAPT
- Classic NAT
- IPsec
- Tunnels
- Content-Aware Proxy
- Traffic-Aware Proxy
- Application Proxy
- Circuit Proxy
- Stateful Packet Filter
- Packet Filter
- Frame Filter
Operational Aspects

• How do you ping/traceroute?
  – DHCP/DNS linkage
  – IPCP linkage
  – Layer 2 information

• What about tunnels?

• What about NAT?
Single point of failure: single-homed routing

ISP
optional NAT
Router

ISP-assigned private address space
Registered address space directly allocated or ISP suballocation

Enterprise

Static Routing or Keepalive
Default Route

ISP-assigned private address space
Simple Multihoming to a Single Provider

Enterprise
Registered address space directly allocated or ISP suballocation
*ISP-assigned private address space*
Simple Multihoming to Two Providers

Enterprise
Registered address space directly allocated or ISP suballocation

Primary ISP POP
Router
More Preferred Default Route

Backup ISP POP
Router
Less Preferred Default Route
RFC 1998 Multihoming

ISP
with 192.0.0.0/16
assigns 192.0.2.0/22 to customer

POP1
advertisises Internet routes

advertisises subset of provider space marked NO-EXPORT

Router 1
iBGP
customer with private AS

numbers west side hosts in 192.0.2.0/23 and east side in 192.0.3.0/23

Router 2
eBGP
Advertises 192.0.2.0/22 192.0.3.0/23

ISP
with 192.0.0.0/16
assigns 192.0.2.0/22 to customer

POP2
dvertisises Internet routes

ISP
with 192.0.0.0/16
assigns 192.0.2.0/22 to customer

POP1
advertisises Internet routes
advertisises Internet routes

ISP
with 192.0.0.0/16
assigns 192.0.2.0/22 to customer

POP2
dvertisises Internet routes

Have I been solving the right problem?

This is L3/IP thinking

Hosted Server 1

ISP A POP 1
optional NAT
Router

Enterprise
Registered address space directly allocated or ISP suballocation
ISP-assigned private address space

ISP B POP 1
optional NAT
Router
Transport/Session

- Load Balancers
- Security gateway (very careful about liability)
  - SSL concentrator
  - IPSec gateway
  - Traffic-inspecting firewall
    - Inappropriate language (problematic)
    - Malware
  - Application layer gateway
- Session Border Controller
  - Principally for VoIP/SIP
  - May be cleaner than firewall for variable ports
Local Distribution

Enterprise
Registered address space directly allocated or ISP suballocation
*ISP-assigned private address space*

ISP POP 1
optional NAT
Router

TCP Load Distributor

ISP POP 1
optional NAT
Router
Global Distribution, Single ISP